Upregulation of c-*MYC* in *cis* through a Large Chromatin Loop Linked to a Cancer Risk-Associated Single-Nucleotide Polymorphism in Colorectal Cancer Cells[▽]

Jason B. Wright,[1] Seth J. Brown,[1] and Michael D. Cole[1,2]*

Department of Genetics[1] and Department of Pharmacology and Toxicology,[2] Dartmouth Medical School, Norris Cotton Cancer Center, Lebanon, New Hampshire 03756

Genome-wide association studies have mapped many single-nucleotide polymorphisms (SNPs) that are linked to cancer risk, but the mechanism by which most SNPs promote cancer remains undefined. The rs6983267 SNP at 8q24 has been associated with many cancers, yet the SNP falls 335 kb from the nearest gene, c-*MYC*. We show that the beta-catenin–TCF4 transcription factor complex binds preferentially to the cancer risk-associated rs6983267(G) allele in colon cancer cells. We also show that the rs6983267 SNP has enhancer-related histone marks and can form a 335-kb chromatin loop to interact with the c-*MYC* promoter. Finally, we show that the SNP has no effect on the efficiency of chromatin looping to the c-*MYC* promoter but that the cancer risk-associated SNP enhances the expression of the linked c-*MYC* allele. Thus, cancer risk is a direct consequence of elevated c-*MYC* expression from increased distal enhancer activity and not from reorganization/ creation of the large chromatin loop. The findings of these studies support a mechanism for intergenic SNPs that can promote cancer through the regulation of distal genes by utilizing preexisting large chromatin loops.

In recent years, genome-wide association studies (GWAS) have provided valuable insights into inherited genetic components of many cancer types. Highly penetrant genes contribute to less than 5% of inherited susceptibility to familial colorectal cancer (CRC) (21). The remaining inherited risk is likely to be accounted for by a combination of many lowly penetrant germ line variations (9, 40). The availability of comprehensive single-nucleotide polymorphism (SNP) arrays allows the large-scale mapping of common genetic variants across the genome and exposes lowly penetrant genetic markers of disease susceptibility. One advantage of SNP-based GWAS is that they identify linkage to susceptibility in both coding and noncoding regions in an unbiased manner. This comprehensive analysis can identify lowly penetrant susceptibility loci beyond specific candidate genes, revealing the potential influence of noncoding regions within the genome.

SNPs within chromosomal region 8q24 have been found to be associated with colorectal, breast, bladder, ovarian, and prostate cancers. Approximately one-third of all reported prostate and colon cancer-associated SNPs map to 8q24 (9, 10, 20). Interestingly, all of the cancer-associated SNPs in 8q24 lie within a 1.5-Mb gene-free region (or "gene desert") bounded by the gene *FAM84B* on the centromeric end and the proto-oncogene c-*MYC* on the telomeric end. The cancer risk-associated SNPs within 8q24 span nearly 600 kb in the 5′ direction from c-*MYC*. Curiously, no SNPs with disease association have been mapped to the even larger gene-free region in the 3′ direction from c-*MYC*. Large gene-free regions are found at diverse locations in the genome, but their significance remains

unclear (31). However, there is ample evidence for distal regulatory influences on c-*MYC* expression. Variant plasmacytoma chromosomal translocation breakpoints were mapped from 150 to 300 kb in the 3′ direction from c-*MYC* (6). Similar chromosomal translocation breakpoint clusters are found in the related human Burkitt's lymphomas (19). These distal translocations misregulate c-*MYC* in *cis* because only the translocated allele is expressed in tumor cells (4). Further evidence for distal regulation of c-*MYC* comes from retrovirally induced T- and B-cell lymphomas in mouse and rat, which frequently target the same distal translocation breakpoint regions (13, 41). A deep body of research identifies c-*MYC* as a hub for signals emanating from innumerable growth and differentiation factors (25), in addition to these distal oncogenic lesions.

Only a small fraction of the human genome encodes proteins, while the remainder consists of 5′ and 3′ flanking regions and introns of various sizes. Dispersed throughout these noncoding regions are evolutionarily conserved sequences that are thought to be potential regulatory elements (31), although few have been experimentally tested. Regulatory elements proximal to promoters have been studied in depth, but the mechanism by which distal elements may regulate gene expression remains unclear. The simplest model is that distal regulatory elements control their target genes through direct recruitment to a promoter, facilitated by chromosome loops. Studies of distal regulatory elements and cellular promoters pose the considerable technical challenge of establishing physical linkage within the cell. The development of the chromosome conformation capture (3C) technique has offered novel insight into large chromatin loops and even potential interchromosomal interactions (reviewed in reference 27). Several gene loci have been shown to exhibit extensive looping to bring regulatory elements into close proximity to their target genes. The full extent of chromosomal looping in gene regulation

* Corresponding author. Mailing address: Department of Genetics, Dartmouth Medical School, Norris Cotton Cancer Center, Lebanon, NH 03756. Phone: (603) 653-9975. Fax: (603) 653-9952. E-mail: mcole @dartmouth.edu.

remains to be explored; however, recent important findings have pointed to very complex genome-wide chromatin organization mediated by loops. Such a model of interconnected chromatin regions may be applied to findings from GWAS to further our understanding of complex regulatory systems involved in disease predisposition (24).

Distinct oncogenic pathways are commonly activated for specific cancers. The majority of CRCs activate the WNT pathway, either through the loss of the adenomatous polyposis coli (APC) tumor suppressor or through oncogenic mutations in beta-catenin (11, 32). Both lesions result in the accumulation of nuclear beta-catenin, which forms a complex with the T-cell factor (TCF)/lymphoid enhancer binding factor 1 (LEF1) transcription factor to activate Wnt target genes (22). A prominent target of the WNT pathway is the c-*MYC* gene, which is an essential downstream effector of the pathway (16, 36, 43). TCF/LEF1 binding sites proximal to c-*MYC* have been found previously (16, 46), but the precise site that functions to enhance c-*MYC* expression has been elusive.

Given the well-known connection of c-*MYC* to colon, prostate, and breast cancers (30), we hypothesized that cancer risk-associated SNPs may alter *cis*-acting regulatory elements that activate c-*MYC* over a great distance. Considerable attention has focused on one SNP, rs6983267, which maps to a TCF4 binding site 330 kb from the c-*MYC* gene, with enhanced TCF4 binding to the cancer risk-associated allele (33, 42). This distal element contains histone modifications that are expected hallmarks of enhancers, with more prominent histone modifications on the cancer-associated allele. We and others have shown that the SNP maps to a region that forms a large chromatin loop, directly linking the SNP to the c-*MYC* promoter (33). Furthermore, we show that the c-*MYC* allele linked in *cis* to the cancer risk-associated SNP variant shows significantly higher expression than the c-*MYC* allele linked to the non-risk-associated variant. Our data support the model that enhanced TCF4 binding to the cancer risk-associated SNP promotes susceptibility to diverse neoplasias by driving elevated c-*MYC* expression in *cis*. Furthermore, we show that the SNP itself does not alter the frequency of looping interactions and that the loop exists in tissues in which there is no known Wnt signaling.

MATERIALS AND METHODS

Evolutionary conservation and binding site prediction. Data on 2 kb of the human genomic sequence surrounding rs6983267 were retrieved from Ensembl and used to create a multispecies alignment of the corresponding regions in rhesus monkey, mouse, dog, horse, armadillo, opossum, and platypus with BLASTN and the Ensembl data-mining tool BioMart (3a). A local alignment of the regions around rs6983267 was prepared using ClustalW and Boxshade (see Fig. 1A). For each species, 25 bp of the aligned sequence surrounding rs6983267 was parsed using the UNIX grep and perl tools (3a). These 25-bp sequences were then analyzed by MEME to discover overrepresented motifs (2, 3). The most statistically significant overrepresented motif found in all sequences was then analyzed using the tom-tom motif comparison tool and the Transfac transcription factor database (44). Tom-tom was used to query the motif against the Transfac database, and known target motifs were ranked by E value. By this method, the highest-scoring position-weighted matrices were found to be dominated by members of the TCF family, indicating a putative TCF binding site over the rs6983267 SNP (see Fig. 1B).

ChIP. Chromatin immunoprecipitation (ChIP) assays were performed using the ChIP assay kit according to the protocol of the manufacturer (Upstate/Millipore). The antibodies used for each respective ChIP were as follows: anti-TCF4 (05-511; Upstate), anti-beta-catenin (06-734; Upstate),

anti-acetylated histone H3 (06-599; Upstate), anti-acetylated histone H4 (06-866; Upstate), and antibody to histone H3 dimethylated at lysine 4 (ab32356; Abcam). PCR amplification of each ChIP or input sample was performed using Ex-*Taq* polymerase (Takara). Labeled and nonlabeled PCR mixtures used 28 and 33 cycles, respectively, and were resolved by gel electrophoresis. Labeled products were quantified using a Typhoon phosphorimager. ChIP products were quantified as a ratio to the input band density by using the corresponding primer pairs. Each ChIP was performed three independent times with HCT116 and DLD-1 CRC cell lines. Error bars in figures indicate standard deviations among results from the three independent ChIPs for each antibody. ChIP primers for rs6983267, rs1447295, rs6942880, and *CCND1* are available upon request.

***In vivo* allele-specific DNA binding.** Allele-specific binding differences were detected directly by sequencing. DLD-1 cells were determined to be heterozygous for the rs6983267(G/T) SNP by sequencing of the PCR product. To normalize variations in the sequencing dye, only sequence profiles from the same run were compared and all samples were amplified in the same number of cycles. G/T allele ratios for each ChIP were normalized with respect to input G/T allele ratios. Products of three independent TCF4 and beta-catenin ChIPs were sequenced, and the G/T allele ratios were averaged. Error bars in the figures indicate the standard deviation of the (ChIP product G/T allele)/(input G/T allele) ratios for each ChIP.

In vivo allele-specific binding differences for TCF4 and beta-catenin were also detected by restriction digestion. ChIP PCR products were digested with the restriction enzyme Tsp45I (NEB). The rs6983267(T) allele generates a Tsp45I site, so allele-specific binding differences can be quantitated by digestion. Equivalent amounts of PCR products were digested, resolved by agarose electrophoresis, and detected with ethidium bromide. The input PCR sample acted as a baseline, and cloned rs6983267(T) and rs6983267(G) were used as restriction enzyme controls. Quantitation of the digestion products by using ethidium bromide band intensity was done with the ImageJ software package (National Institutes of Health). Densitometry analysis was done using the analysis module, and the quantitation was normalized with respect to input controls.

3C. 3C was used to test for intrachromosomal looping interactions. We performed cross-linking, digestion, and ligation using established procedures (8, 28, 39). Briefly, cells (10^8 per experiment) were cross-linked with 1% formaldehyde, which was then quenched with 0.125 M glycine. Cell pellets were homogenized in lysis buffer (10 mM Tris-Cl, pH 8.0, 10 mM NaCl, 0.2% [vol/vol] NP-40) with a Dounce homogenizer B. The cell pellet was split into 20 aliquots of approximately 5×10^6 cells each. To digest cross-linked cells, the cell pellets were placed in the appropriate restriction enzyme buffer and cells were incubated in 0.3% sodium dodecyl sulfate (SDS). Triton X-100 (1%) was added to each tube to sequester SDS. Cells were digested with HindIII, SacI, and PstI restriction enzymes in separate experiments with 400 U per 5×10^6 cells for 36 h at 37°C. One aliquot of digested cells was set aside to test restriction enzyme efficiency by electrophoresis and PCR. Each aliquot was ligated with 4,000 U of T4 DNA ligase for 2 h at 16°C. Cross-links were reversed by proteinase K digestion overnight, and the 3C DNA template was purified by two rounds of phenol-chloroform extraction followed by five rounds of ethanol precipitation. The final DNA pellet was dissolved in Tris-EDTA buffer, pH 8.0, and quantitated by using a NanoDrop spectrophotometer.

To control for PCR and primer efficiencies, PCR products were first made for all primer pairs spanning 8q24 (data not shown). Primers were chosen within the region from bp 80 to −160 relative to the target restriction site, and the annealing temperature was uniformly 58 ± 1°C. The amplified products from the restriction sites were pooled, mixed with genomic DNA, digested, and ligated. Ligation products were then amplified with one primer distal to c-*MYC* and one primer proximal to c-*MYC* to assess the efficiency of cross ligation (data not shown). To verify that the cross-ligation products were as expected, the amplified products were cloned into pGemT and sequenced. All PCRs were performed in triplicate, and each set of error bars in the figures indicates the standard error of the mean represented by each data point.

Cell culture and allele-specific expression. DLD-1 and K562 cells were maintained in RPMI 1640 with 10% serum and 1% penicillin-streptomycin. MCF7 and HCT116 cells were maintained in Dulbecco's modified Eagle medium with 10% serum and 1% penicillin-streptomycin. Polymorphism screens across the c-*MYC* locus were run using primers corresponding to overlapping 500-bp PCR segments spanning the c-*MYC* transcribed region and 2 kb of the proximal promoter (primers will be disclosed upon request). Nuclear RNA was isolated from DLD-1 cells by using Trizol and ethanol precipitated. Expression was evaluated by reverse transcription PCR (RT-PCR) using primers flanking the rs4645953 polymorphism within the first intron of c-*MYC*
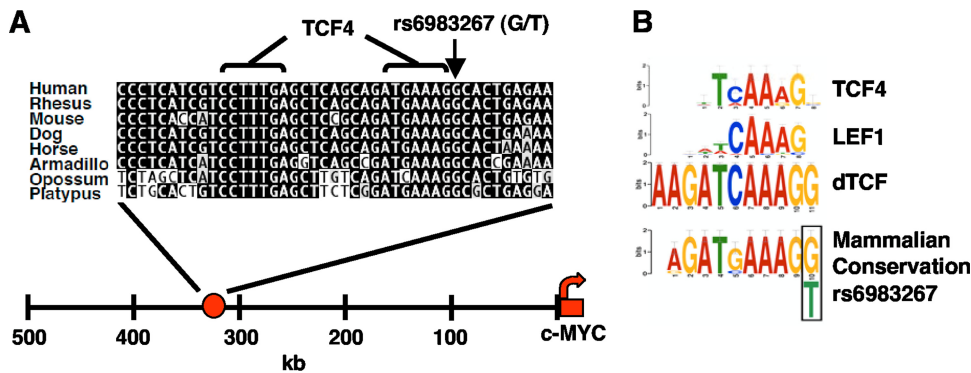
FIG. 1. The rs6983267 SNP is highly conserved in mammalian evolution and maps to a consensus TCF/LEF1 binding site. (A) Comparative genomics analysis of the rs6983267 SNP regions in the indicated mammals, including marsupials and monotremes. The ancestral element has a G, whereas the allelic frequency in humans is 50% T. Consensus inverted TCF binding sites are indicated by brackets. rs6983267 falls ~335 kb in the 5′ direction from c-*MYC* in all mammals. (B) Alignment of the three top-scoring consensus TCF/LEF1 motifs (dTCF, *Drosophila* T-cell factor) is shown above the mammalian conserved sequence surrounding the rs6983267 SNP. Both the ancestral G variant, which is strongly associated with several cancers, and the non-cancer-associated T variant of rs6983267 are depicted (the variations are indicated by the rectangular black box).

in DLD-1 cells. Allelic expression ratios were determined by using sequencing dye trace profile peaks at the site of the rs4645953 SNP. To control for slight dye imbalance within the sequencing runs, PCR products from genomic DNA of the corresponding region surrounding the rs4645953 SNP were sequenced in tandem with each expression profile. Quantitation of allele-specific profiles was carried out by measuring the C/T allele ratio relative to the genomic DNA C/T allele ratio. To control for DNA contamination, each RT-PCR was controlled with a no-RT PCR (data not shown). Additionally, RNase digestion prior to each RT-PCR confirmed that each product was amplified from an RNA template.

To establish a control for allele-specific expression, we used a cloned human c-*MYC* gene containing the native c-*MYC* promoter and 7 kb of the 5′ flanking sequence. Site-directed mutagenesis was used to create matched expression vectors that differed only at the rs4645953 SNP. The cloned polymorphic c-*MYC* genes (4 μg total DNA) were mixed at either a 1:1 or a 2:1 ratio, and 10^6 Rat1a fibroblast cells were transfected with the mixtures by using Fugene transfection reagents. RNA was isolated by using Trizol, and exogenous c-*MYC* expression was detected using human-specific primers. Allelic expression was evaluated as described above.

RESULTS

We are interested in defining the mechanism through which individual genomic SNPs contribute to cancer susceptibility. The rs6983267(G/T) SNP lies approximately 335 kb in the 5′ direction from the c-*MYC* gene and has allele frequencies of 0.5. The G variant correlates with predisposition to colorectal, prostate, and ovarian cancers with odds ratios of 1.2 to 1.3 for heterozygotes and 1.4 to 1.5 for homozygotes (1, 9, 10, 15, 20, 40). Notably, the rs6983267 SNP lies within a 1.3-kb region of 8q24 that is strongly conserved among mammals (Fig. 1A). The rs6983267(G) allele correlates with mammalian conservation, indicating that it is the ancestral allele. Synteny of this element with c-*MYC* is maintained in all mammals, suggestive of a conserved *cis*-regulatory role. To explore the function of rs6983267, we noted that the SNP itself lies within a consensus binding site for the transcription factor TCF/LEF1, a known downstream effector of the WNT pathway and an established activator of the c-*MYC* gene (Fig. 1B) (16, 22). The polymorphism lies at a position within the TCF motif known to affect binding affinity, favoring the G variant over the T variant (14). In fact, this conserved chromosomal element contains two inverted TCF/LEF1 consensus sites, with the polymorphism at

the outer edge of the rightward site (Fig. 1A). Based on binding site prediction, we hypothesized that the respective rs6983267 SNP variants (the G and T variants) would differentially bind to TCF4–beta-catenin in human CRC cells.

Enhanced beta-catenin–TCF binding to the cancer risk-associated rs6983267 allele. To examine binding of the rs6983267 SNP to TCF *in vivo*, we analyzed two different CRC cell lines: HCT116, which bears a mutation in beta-catenin rendering the beta-catenin stable, and DLD-1, which has an inactivating mutation in the APC gene (29). Both lesions provide a constitutively active WNT signaling pathway. Using ChIP, we found that binding of both TCF4 and beta-catenin is strongly enriched in a 300-base region centered on the rs6983267 SNP in both of the CRC cell lines (Fig. 2A). Neither factor is detected at rs1447295 or rs9642880, nearby SNPs which show no association with colon cancer (Fig. 2A). Binding to a known TCF4 site in the *CCND1* gene served as a positive control (38). This indicates that TCF4 and beta-catenin bind at or near rs6983267 *in vivo*.

Fortuitously, DLD-1 cells are diploid for chromosome 8 and heterozygous for the rs6983267(G/T) SNP variants, which made it possible to compare TCF4–beta-catenin binding affinities between the alleles *in vivo*. We compared the rs6983267 allele distribution in the input DNA to the distribution after immunoprecipitation with either beta-catenin or TCF4 antibodies. Differences in allelic binding at rs6983267 were resolved by direct sequencing of the PCR products. The input sequence shows equal contributions from the alleles, as expected in a heterozygous DNA sample (Fig. 2B). Compared to input DNA, the TCF4 ChIP sample shows threefold enrichment with the rs6983267(G) allele and the beta-catenin ChIP sample shows fourfold enrichment with the rs6983267(G) allele over the rs6983267(T) allele (Fig. 2D). As an alternate method to assess allele-specific binding, the input and ChIP DNA samples were digested with the restriction enzyme Tsp45I. The rs6983267(T) allele has a cleavage site for this enzyme that is not present in the rs6983267(G) allele. The input DNA samples show approximately equal G and T allele distributions (Fig. 2C), consistent with a heterozygous geno-
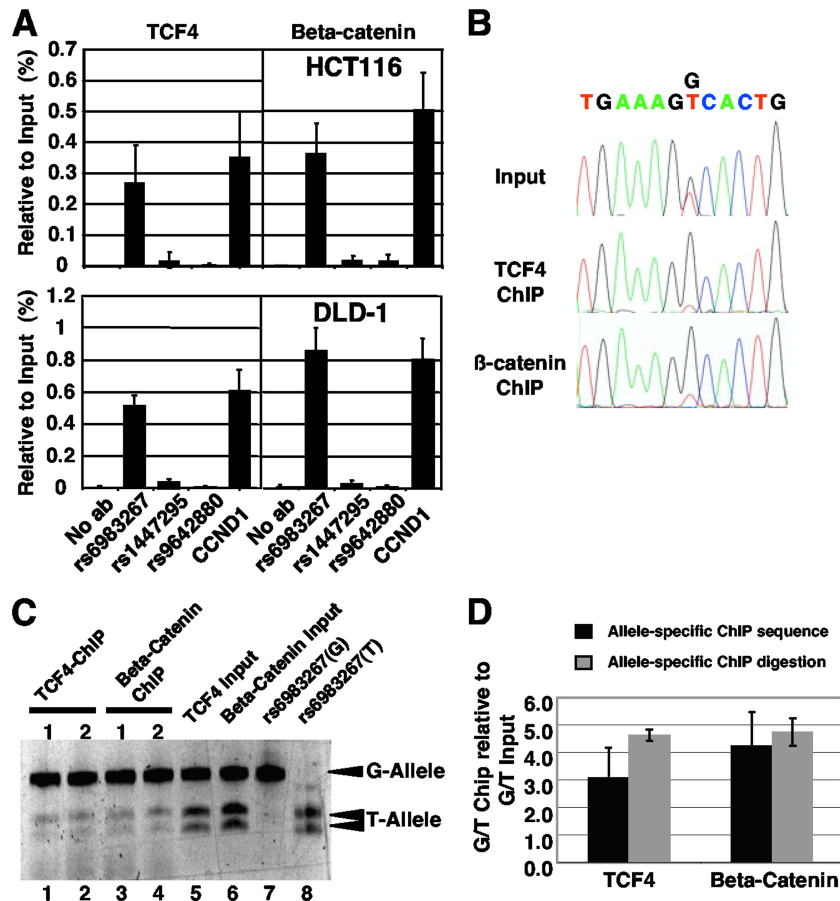
FIG. 2. Enhanced binding of TCF4 and beta-catenin to the cancer-associated rs6983267 SNP *in vivo*. (A) ChIP analysis of TCF4 and beta-catenin binding in CRC cell lines HCT116 and DLD-1. Samples from independent, duplicate immunoprecipitations were assayed for the presence of three different cancer-associated SNPs: rs6983267 (associated with colon cancer and others), rs1447295 (associated with prostate cancer), and rs6942880 (associated with bladder cancer). A TCF4 binding site near the *CCND1* gene was used as a positive control. ChIP allowed quantitation of TCF4 and beta-catenin binding to rs6983267 and other SNPs. The ChIP signal (expressed as a percentage relative to the input level) is shown for three different SNPs, as well as for the control TCF4 site in the *CCND1* gene. No ab, no antibody. (B) The rs6983267 SNP region from DLD-1 cells was PCR amplified from either input DNA or chromatin immunoprecipitates obtained using anti-TCF or anti-beta-catenin antibodies. The PCR products were subjected to direct DNA sequencing using a primer from one end. The input DNA is heterozygous for rs6983267, whereas the ChIP DNA samples obtained with both TCF4 and beta-catenin antibodies show enrichment with the cancer-associated rs6983267(G) allele. (C) Restriction enzyme analysis of allele distribution in DLD-1 input DNA compared to that in ChIP DNA by using anti-TCF4 and anti-beta-catenin. Tsp45I cleaves the rs6983267(T) allele but not the rs6983267(G) allele. rs6983267(T)-only and rs6983267(G)-only amplification products were used to verify specificity and complete enzyme digestion (lanes 7 and 8). Independent duplicate ChIP samples were amplified for each antibody (lanes 1 to 4), and equivalent amounts of DNA were digested and loaded into the lanes. Equivalent samples of input DNA for the individual immunoprecipitations were also digested and loaded (lanes 5 and 6). (D) Quantitation of rs6983267 allelic distributions in input and ChIP samples by either direct sequencing or Tsp45I digestion. For each antibody, the peak height of the G/T allele ratio for the input DNA was set to 1 and used to normalize the G/T allele ratio for the ChIP DNA. All DNA samples were amplified in the same number of cycles and to the minimum extent needed for direct sequencing or digestion.

type and the sequencing profiles described above. The ChIP DNA samples obtained with TCF4 and beta-catenin antibodies both show approximately fourfold enrichment with the rs6983267(G) allele (Fig. 2D), consistent with the sequencing profiles. Together, these data show that TCF4 and beta-catenin preferentially occupy the rs6983267(G) SNP variant over the rs6983267(T) variant *in vivo*. These data are consistent with recent findings (33, 42).

Enhancer-specific chromatin modification at the rs6983267 SNP. To further test the hypothesis that the rs6983267 SNP could regulate c-*MYC*, we wanted to explore if the SNP falls within an enhancer-like element. Enhancers are capable of regulating transcription from distal locations relative to their

respective target genes, and recent ChIP studies have uncovered distinguishing histone modifications and cofactors that delineate enhancers and promoters (17, 18, 26). To this end, we used ChIP to analyze several histone modifications (dimethylation of histone H3 at lysine 4 [H3K4me2], acetylation of histone H3 [H3-Ac], and H4-Ac) which are hallmarks of chromosomal enhancers. ChIP assays for all three enhancer marks showed enrichment at the site of the rs6983267 SNP (Fig. 3A). Nearby prostate cancer SNPs (rs1447295 and rs9642880) show no enrichment with these histone marks, indicating the specificity of this enhancer-predictive signature in human CRC cells (Fig. 3A). A known TCF binding site in the *CCND1* gene displays the same histone modification profile and served as a
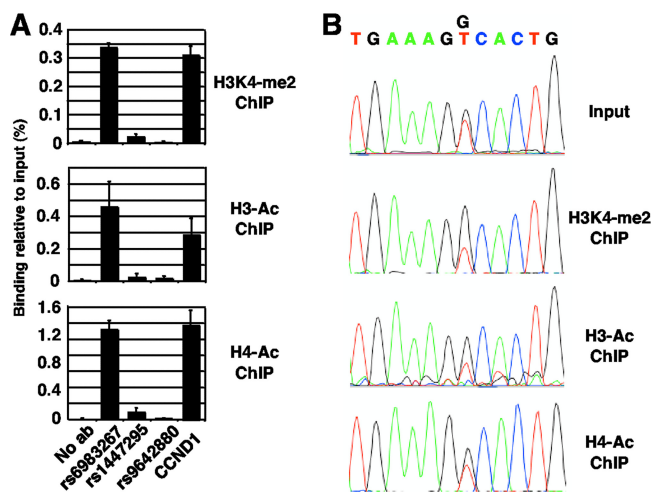
FIG. 3. Allele-specific enhancer histone modifications at the cancer-associated rs6983267 SNP. (A) Quantitation of three different enhancer-associated histone modifications (H3K4me2, H3-Ac, and H4-Ac) at rs6983267, rs1447295, and rs6942880, as well as the region surrounding the TCF4 binding site in *CCND1*, by ChIP. Data are presented as percentages of the input level. (B) Allele specificity of histone modifications as determined by direct sequencing of the PCR products.

positive control. To further explore the allele specificity, we sequenced the PCR product from the histone ChIP and found that the rs6983267(G) variant was enhanced with H3K4me2, H3-Ac, and H4-Ac modifications (Fig. 3). Thus, the rs6983267 SNP region has the hallmarks of a chromosomal enhancer, and they are more prominent for the cancer-associated rs6983267(G) allele.

A large chromatin loop links rs6983267 to the c-*MYC* promoter. Given the long chromosomal distance between rs6983267 and the c-*MYC* promoter, we pursued a looping mechanism of enhancer recruitment. To explore this possibility, we used 3C, a technique that can show functional connectivity between enhancer or insulator elements and their respective target genes through chromatin loops (8, 39). DLD-1 CRC cells were cross-linked with formaldehyde, and the chromatin was digested with HindIII and ligated. Intrachromosomal interactions were assayed using primer pairs arrayed at intervals across a 350-kb segment of 8q24 spanning rs6983267, several other cancer-associated SNPs, and the c-*MYC* promoter. A more extensive set of primers was used for HCT116 cells, with primers spaced at approximately 18-kb intervals. Remarkably, the 3C analysis shows a clear peak of interaction with the c-*MYC* promoter at the site closest to the rs6983267 SNP in both DLD-1 and HCT116 cells (Fig. 4A and B). Ligation of distal chromosome segments was verified by direct sequencing, validating the distal interaction. The two primers flanking the HindIII site near rs6983267 scored positive for looping with equal efficiencies by PCR (data not shown). The large chromatin loop between rs6983267 and the c-*MYC* promoter was detected in three independent cross-linked cell populations for each cell line. Potentially weaker interactions with the c-*MYC* promoter were observed with primer E (15 kb from rs6983267) and primer R (6 kb from the c-*MYC* promoter) (Fig. 4A and B). We do not know if these interactions represent indepen-

dent chromatin loops or local chromatin interactions that are expected with the use of this technique (7). Notably, two other cancer-associated SNPs which predispose to prostate or bladder cancer show no evidence of looping in colon cancer cells (Fig. 4A and B). Results from control experiments show that all possible PCR and ligation products could have been efficiently detected if present (see Materials and Methods). To further validate the findings from the 3C studies, we designed a second set of seven primer pairs centered on PstI restriction sites instead of HindIII sites. Only the site nearest the rs6983267 SNP scored positive for a chromatin loop with the c-*MYC* promoter (data not shown). Thus, we conclude that the rs6983267 SNP forms a 335-kb chromatin loop and interacts directly with the c-*MYC* promoter in two different CRC cell lines, with maximal interaction seen proximal to the transcriptional start site. These data are consistent with recent findings (33).

To explore any potential tissue specificity of the chromatin loop, we conducted 3C experiments with two additional cell lines, MCF7 breast cancer cells and K562 myeloid leukemia cells. We used the same cross-linking, digestion, and ligation conditions used for the CRC cells and assayed for chromatin loops by PCR using an extensive set of primers spanning 8q24. As in CRC cells, we detected a robust 3C signal using primers from rs6983267 and c-*MYC* (Fig. 4C and D). No other 3C ligations were detected with other primers between rs6983267 and c-*MYC*, except for a local interaction with a primer proximal to c-*MYC* itself. These data suggest that the rs6983267-c-*MYC* chromatin loop may be present in all developmental lineages.

To explore if the rs6983267 enhancer interacts with other promoters, we expanded the 3C studies to compare its interaction with c-*MYC* to interactions with two other promoters in the region, POU5FP1 and PVT1. POU5FP1 is identified as a pseudogene in GenBank, whereas PVT1 is a noncoding RNA. We detect a weak interaction between rs6983267 and the POU5FP1 promoter, which is only 15 kb away. We detect a more significant interaction between the rs6983267 enhancer and the PVT1 promoter, which maps 393 kb away and 58 kb in the 3′ direction from the c-*MYC* promoter. This finding is consistent with the observation that the rs6983267(G) allele is correlated with PVT1 expression (34). However, the interaction between the rs6983267 enhancer and the c-*MYC* promoter was substantially more robust. We detect no interaction between the rs6983267 enhancer and the prostate cancer risk-associated SNP rs1447295.

Allele-independent chromatin looping. We were interested to further analyze the potential influence of the rs6983267 SNP on the chromatin loop. Another primer from rs6983267 was designed to incorporate the SNP into the 3C PCR products from the polymorphic DLD-1 cells, which were then sequenced directly as described above for the ChIP samples. The sequence profiles indicate that the cancer risk- and non-risk-associated rs6983267 alleles form loops with the c-*MYC* promoter with equal efficiencies (Fig. 5). In support of this conclusion, we conducted a 3C experiment after cross-linking MCF7 breast cancer cells, which are also heterozygous for the rs6983267 SNP but are not known to have exaggerated beta-catenin–TCF signaling as in CRC cells. The two alleles formed
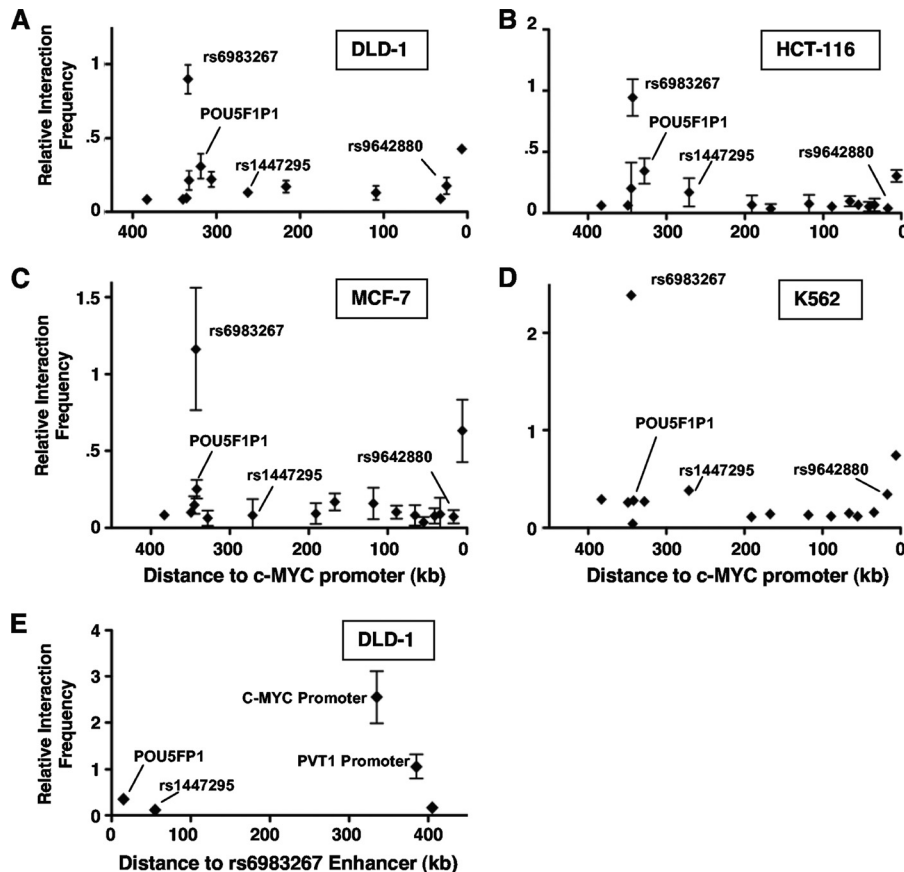
FIG. 4. A 335-kb chromatin loop links the cancer-associated rs6983267 SNP to the c-*MYC* promoter. (A to D) 3C analyses of the c-*MYC* promoters in DLD-1 CRC cells, HCT116 CRC cells, MCF7 breast cancer cells, and K562 myeloid leukemia cells. Relative frequencies of interaction between primers derived from the 5′ flanking domain and the c-*MYC* promoter are shown. The data in panels A, B, and C are derived from three separate 3C analyses for each primer combination, and those in panel D are from one experiment. PCR products were normalized relative to positive-control templates. An interaction of the c-*MYC* promoter with rs6983267 but not with two other cancer-associated SNPs or other primers in the region was observed. (E) 3C analysis of the rs6983267 enhancer in DLD-1 cells. A primer from rs6983267 was combined with primers from POU5FP1, rs1447295, c-*MYC*, or PVT1. Relative interaction frequencies were calculated as described above.

rs6983267-c-*MYC* chromatin loops with equal frequencies, as did those in DLD-1 cells (Fig. 5). These data suggest that the rs6983267-c-*MYC* chromatin loop is mediated by factors other than beta-catenin–TCF binding.
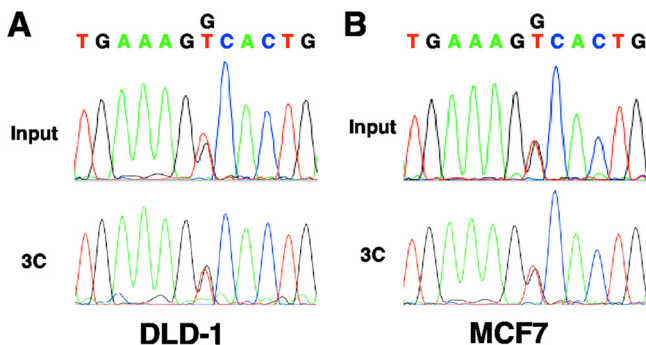

FIG. 5. Chromatin looping is not influenced by the rs6983267 SNP. The allele specificity of intrachromosomal looping was evaluated by sequencing of the 3C PCR product from an rs6983267/c-*MYC* primer pair. The two rs6983267 alleles are detected equally in both DLD-1 colon cancer cells and MCF7 breast cancer cells.

Enhanced expression of the cancer risk-associated c-*MYC* allele. Recent studies have shown that, in an exogenous system, the rs6983267(G) variant can drive *in vitro* luciferase expression in response to WNT signaling more effectively than rs6983267(T), supportive of its role as an SNP-sensitive TCF4–beta-catenin enhancer (33). However, there has been no direct evidence that links individual rs6983267 alleles to the level of c-*MYC* expression in CRC (34). We reasoned that the difficulty in addressing this issue lies in the complex and variable genetic and cellular backgrounds inherent in a collection of pathology specimens or cell lines, which have various impacts on c-*MYC* expression. Given that the cancer risk-associated SNP in question is present in approximately 50% of the global population and that c-*MYC* expression is tightly regulated, any difference in expression may be subtle and difficult to expose. Therefore, we focused on the relative expression levels of the two different c-*MYC* alleles within a single cell line, the DLD-1 CRC cell line. We hypothesized that if the cancer risk allele was driving c-*MYC* expression in *cis*, then we should be able to detect an imbalance in expression between the two c-*MYC* alleles in cells heterozygous for rs6983267, thus allowing us to compare the effects of two SNP genotypes without introducing additional variables by comparing different cell lines.
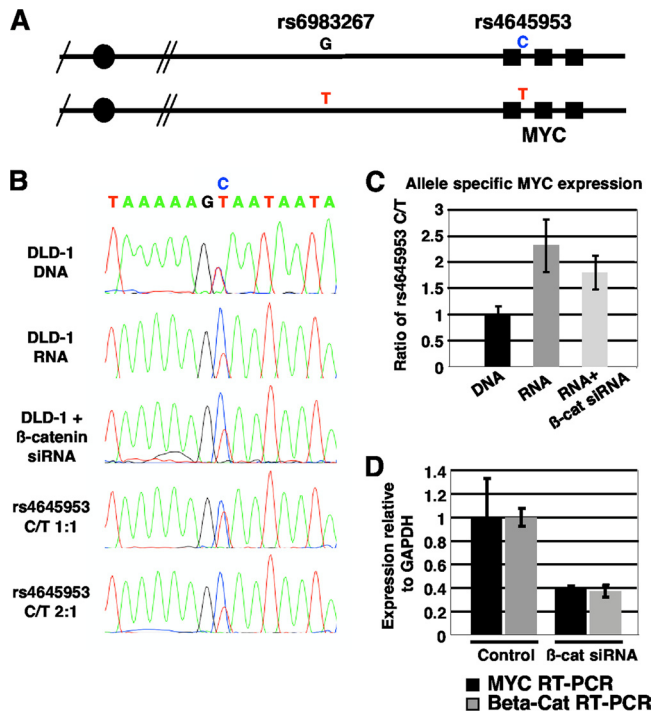
FIG. 6. The cancer risk-associated rs6983267(G) allele is linked to enhanced c-*MYC* expression. (A) Diagram of the linked polymorphisms at c-*MYC* SNPs rs6983267 and rs4645953 in DLD-1 cells as determined by 3C analysis. (B) Allele-specific c-*MYC* expression as evaluated by PCR analysis of the c-*MYC* rs4645953 SNP in DLD-1 cells. Genomic DNA exhibits heterozygosity at the rs4645953 SNP (DLD-1 DNA profile). RT-PCR detects twofold-enhanced expression of the c-*MYC* rs4645953(C) allele (DLD-1 RNA profile). RT-PCR detects the same twofold-enhanced expression of the c-*MYC* rs4645953(C) allele after knockdown of beta-catenin expression with siRNA (DLD-1 + β-catenin siRNA profile). Sequencing profiles accurately detect allele-specific expression because transient transfection with cloned full-length c-*MYC* genes carrying the different rs4645953 alleles (C and T alleles) gives a 1:1 ratio of expression levels when the c-*MYC* genes are mixed in a 1:1 ratio (rs4645953 C/T 1:1 profile). Mixing the c-*MYC* rs4645953 C and T alleles in a 2:1 ratio gives a corresponding 2:1 ratio of expression levels as determined by RT-PCR (rs4645953 C/T 2:1 profile). (C) RT-PCR quantitation of c-*MYC* allele-specific expression in DLD-1 cells before and after knockdown of beta-catenin (β-cat) expression by using siRNA. (D) Knockdown of beta-catenin expression gives a corresponding reduction in c-*MYC* expression. GAPDH, glyceraldehyde-3-phosphate dehydrogenase.

Critical for this experiment is the ability to measure allele-specific expression. Therefore, we screened the c-*MYC* transcribed region in DLD-1 cells for polymorphisms and identified a heterozygosity in the first intron (rs4645953) (Fig. 6A). Extensive sequencing of the locus confirmed that, apart from the rs4645953 polymorphism, there was no heterozygosity between the two c-*MYC* alleles in DLD-1 cells.

If indeed the rs6983267 SNP alters the activity of a distal beta-catenin–TCF4 enhancer, then the c-*MYC* allele in *cis* to the cancer risk-associated allele would show higher expression than the c-*MYC* allele linked to the non-risk-associated SNP. We first had to establish the linkage between the different SNPs at rs6983267 and c-*MYC* (rs4645953), which are separated by 330 kb. To determine which c-*MYC* allele was linked to the cancer risk

rs6983267(G) allele, we repeated the 3C experiment with a different restriction enzyme which allowed the use of primers that span both the rs6983267 and c-*MYC* (rs4645953) SNPs. We again detected a robust 3C ligation product, confirming the presence of the chromatin loop (data not shown). Sequencing of multiple individually cloned 3C PCR products showed linkage between the cancer risk rs6983267(G) allele and the rs4645953(C) c-*MYC* allele (Fig. 6A). Other clones showed linkage of the rs6983267(T) and rs4645953(T) alleles. No cloned 3C PCR products with G-T or T-C allele linkage were found, establishing unambiguous linkage of G-C and T-T alleles on the respective chromosomes.

To study the relative levels of expression of the different c-*MYC* alleles, we performed RT-PCR using primers that bracket the rs4645953 SNP. Direct sequencing of the RT-PCR product provided information on the relative contribution of each allele to c-*MYC* expression. These data show that the c-*MYC* rs4645953(C) allele was expressed at a level approximately twofold higher than the c-*MYC* rs4645953(T) allele (Fig. 6B and C), implying that the linked rs6983267(G) SNP results in significant enhancement. This ratio is significant because the two c-*MYC* genes are expressed equally in most cells (23) and most genes in the genome exhibit biallelic expression (12). We have confirmed equal levels of expression of the two c-*MYC* alleles in mouse fibroblasts with a c-*MYC* polymorphism (data not shown). To validate that the sequencing profiles from the RT-PCR experiment give an accurate representation of relative c-*MYC* allele expression levels, we created cloned full-length c-*MYC* genes driven by the native c-*MYC* promoter that differed only at the rs4645953(C/T) SNP. Rat1a cells were then transiently transfected with these genes at a DNA (C/T allele) ratio of either 1:1 or 2:1, and expression was assayed by RT-PCR as described for the DLD-1 CRC cells. Exogenous expression of the two constructs showed equivalent allelic expression levels at a 1:1 DNA transfection ratio and a corresponding 2:1 ratio of expression levels at the 2:1 DNA transfection ratio (Fig. 6B). The results of these control experiments show that the rs4645953 SNP does not itself alter c-*MYC* expression and that the sequencing profiles can accurately detect the differential expression of different alleles within a cell.

To extend these expression studies further, we asked if allele-specific c-*MYC* expression is dependent on beta-catenin. The level of beta-catenin was knocked down in DLD-1 CRC cells by using small interfering RNA (siRNA), and c-*MYC* expression was monitored by RT-PCR. The beta-catenin siRNA led to a 60% reduction in c-*MYC* expression (Fig. 6D), which is consistent with previous data obtained using a dominant negative form of TCF (43). Interestingly, the c-*MYC* rs4645953(C) allele continued to show a twofold-higher level of expression than the c-*MYC* rs4645953(T) allele (Fig. 6B), despite the net reduction in total c-*MYC* mRNA.

DISCUSSION

Numerous GWAS link specific SNPs to cancer, but the mechanism by which they contribute to cancer risk has remained elusive. This is especially true for SNPs that are distal to any potential effector gene, such as the many cancer risk SNPs mapped to 8q24. We show in this study that a SNP that is 330 kb away from the c-*MYC* gene can lead to a twofold

change in c-*MYC* expression and that higher-level expression is linked to the cancer risk allele. It is well known that elevated c-*MYC* levels contribute to cancer through somatic mutations such as chromosomal translocations and gene amplification (5). Elevated c-*MYC* levels also arise through oncogenic signaling pathways such as the WNT/APC pathway which activate c-*MYC* transcription (16, 43), and c-*MYC* is an essential effector of the WNT/APC pathway in intestinal cancer (36). Notably, c-*MYC* is haploinsufficient for tumor formation after the loss of APC (45), and even subtle variations in c-*MYC* levels can alter the transformed phenotype in other systems (37). Thus, a twofold elevation in c-*MYC* expression can readily explain the increased cancer risk associated with the rs6983267(G) allele.

It was interesting to find that the dominant rs6983267(G)-linked c-*MYC* expression was sustained after beta-catenin knockdown. One interpretation of this finding is that the rs6983267 enhancer is the major WNT-responsive enhancer driving c-*MYC* expression in colon cancer cells. Reducing beta-catenin levels would reduce the expression of both c-*MYC* alleles, but the rs6983267(G) allele would continue to be dominantly expressed because of its higher affinity for beta-catenin–TCF4. This finding is supported by recent ChIP-chip and ChIP-sequencing analyses of TCF4 complex occupancy across the 8q24 locus in CRC cells, which have highlighted rs6983267 as the strongest TCF4 binding site in the broad c-*MYC* domain (33, 42). Furthermore, the rs6983267 enhancer can recapitulate an expression pattern similar to that of endogenous c-*MYC* in transgenic mice (42).

The allele-specific binding of beta-catenin–TCF4 to the rs6983267 SNP and a direct loop with the c-*MYC* promoter argue strongly that the rs6983267 SNP is a distal enhancer that is directly responsive to the WNT pathway. The large distance between rs6983267 and c-*MYC* greatly reduces the chance of a false positive due to a spurious ligation product in the 3C analysis. Chromatin loops have been detected at several loci (reviewed in reference 27) but rarely over this great a distance. The presence of a regulatory element within the >1-Mb gene desert in the 5′ direction from c-*MYC*, which is otherwise devoid of functional genes, raises the possibility that this domain harbors other regulatory elements which respond to diverse signaling pathways. It is notable that 16 different cancer-associated SNPs have been mapped to this large 5′ flanking domain of c-*MYC* (9, 10, 20, 35). Regulatory elements within this domain may form multiple distinct chromatin loops to facilitate c-*MYC* regulation in different cellular environments. Such regulatory elements are subject to subtle but significant variation by nearby SNPs.

The use of heterozygotic cell lines for analysis of allele-specific regulation has a number of advantages over experimental comparisons between 3C data from different cell lines. Since DLD-1 cells contain both G and T variants of the rs6983267 SNP, the alleles can be compared to each other within a defined cell background, thus bypassing the variability inherent in independently derived cell lines. In comparisons of different 3C libraries, multiple technical variables such as cross-linking conditions and digestion and ligation efficiencies may make subtle differences in loop frequency difficult to discern. We applied the 3C protocol to the rs6983267-c-*MYC* loop and then performed direct sequencing, as was done with the allele-specific ChIP samples. Somewhat surprisingly, our find-

ings indicate that the risk- and non-risk-associated rs6983267 alleles form loops with the c-*MYC* promoter with equal efficiencies (Fig. 5). Equivalent frequencies of loop formation by the different rs6983267 alleles were confirmed by Tsp45I digestion (data not shown). This finding suggests that the loop is present independent of the affinity of the enhancer for beta-catenin–TCF4. We also conducted 3C experiments with MCF7 breast cancer cells and K562 myeloid leukemia cells. Both of these cell lines are heterozygous for rs6983267 but do not have known WNT-dependent regulation of c-*MYC* as in CRC cells. The rs6983267-c-*MYC* chromatin loop was readily detected in both MCF7 and K562 cells, and as in DLD-1 cells, the G and T alleles formed loops with equal frequencies (Fig. 5). The data support a model in which the loop itself is not responsive to active WNT signaling and is present independent of the genotype of the cancer-associated SNP. These data suggest that the rs6983267-c-*MYC* chromatin loop is mediated by factors other than beta-catenin–TCF binding and that predisposition to CRC is driven by *cis*-acting transcriptional activation of c-*MYC* through increased TCF4 recruitment and not by increased loop formation. This contrasts with a recent finding that a more proximal loop with the c-*MYC* promoter can be induced by WNT signaling (47).

It will be interesting to study the rs6983267 SNP in more detail to determine if c-*MYC* exhibits allele-specific expression in normal development or if the impact of this distal enhancer is apparent only in cells with exaggerated WNT signaling, such as CRC cells. Nevertheless, the broad implications of the findings of our study provide a molecular mechanism for the cancer risk-associated rs6983267 to enhance c-*MYC* expression in CRC and also identify a novel regulatory element for the c-*MYC* oncogene. Together, these findings may act as a paradigm for cancer-associated SNPs within 8q24 and other intergenic regions.

REFERENCES

1. **Amundadottir, L. T., P. Sulem, J. Gudmundsson, A. Helgason, A. Baker, B. A. Agnarsson, A. Sigurdsson, K. R. Benediktsdottir, J. B. Cazier, J. Sainz, M. Jakobsdottir, J. Kostic, D. N. Magnusdottir, S. Ghosh, K. Agnarsson, B. Birgisdottir, L. Le Roux, A. Olafsdottir, T. Blondal, M. Andresdottir, O. S. Gretarsdottir, J. T. Bergthorsson, D. Gudbjartsson, A. Gylfason, G. Thorleifsson, A. Manolescu, K. Kristjansson, G. Geirsson, H. Isaksson, J. Douglas, J. E. Johansson, K. Balter, F. Wiklund, J. E. Montie, X. Yu, B. K. Suarez, C. Ober, K. A. Cooney, H. Gronberg, W. J. Catalona, G. V. Einarsson, R. B. Barkardottir, J. R. Gulcher, A. Kong, U. Thorsteinsdottir, and K. Stefansson.** 2006. A common variant associated with prostate cancer in European and African populations. Nat. Genet. **38:**652–658.
2. **Bailey, T. L., and C. Elkan.** 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc. Int. Conf. Intell. Syst. Mol. Biol. **2:**28–36.
3. **Bailey, T. L., N. Williams, C. Misleh, and W. W. Li.** 2006. MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Res. **34:**W369–W373.
3a. **Brown, S. J., M. D. Cole, and A. J. Erives.** 2008. Evolution of the holozoan ribosome biogenesis regulon. BMC Genomics **9:**442.
4. **Cesarman, E., R. Dalla-Favera, D. Bentley, and M. Groudine.** 1987. Mutations in the first exon are associated with altered transcription of c-myc in Burkitt lymphoma. Science **238:**1272–1275.
5. **Cole, M. D.** 1986. The myc oncogene: its role in transformation and differentiation. Annu. Rev. Genet. **20:**361–385.
6. **Cory, S., M. Graham, E. Webb, L. Corcoran, and J. M. Adams.** 1985. Variant (6;15) translocations in murine plasmacytomas involve a chromosome 15 locus at least 72 kb from the c-myc oncogene. EMBO J. **4:**675–681.

7. **Dekker, J.** 2006. The three 'C' s of chromosome conformation capture: controls, controls, controls. Nat. Methods **3:**17–21.

8. **Dekker, J., K. Rippe, M. Dekker, and N. Kleckner.** 2002. Capturing chromosome conformation. Science **295:**1306–1311.

9. **Easton, D. F., K. A. Pooley, A. M. Dunning, P. D. Pharoah, D. Thompson, D. G. Ballinger, J. P. Struewing, J. Morrison, H. Field, R. Luben, N. Wareham, S. Ahmed, C. S. Healey, R. Bowman, K. B. Meyer, C. A. Haiman, L. K. Kolonel, B. E. Henderson, L. Le Marchand, P. Brennan, S. Sangrajrang, V. Gaborieau, F. Odefrey, C. Y. Shen, P. E. Wu, H. C. Wang, D. Eccles, D. G. Evans, J. Peto, O. Fletcher, N. Johnson, S. Seal, M. R. Stratton, N. Rahman, G. Chenevix-Trench, S. E. Bojesen, B. G. Nordestgaard, C. K. Axelsson, M. Garcia-Closas, L. Brinton, S. Chanock, J. Lissowska, B. Peplonska, H. Nevanlinna, R. Fagerholm, H. Eerola, D. Kang, K. Y. Yoo, D. Y. Noh, S. H. Ahn, D. J. Hunter, S. E. Hankinson, D. G. Cox, P. Hall, S. Wedren, J. Liu, Y. L. Low, N. Bogdanova, P. Schurmann, T. Dork, R. A. Tollenaar, C. E. Jacobi, P. Devilee, J. G. Klijn, A. J. Sigurdson, M. M. Doody, B. H. Alexander, J. Zhang, A. Cox, I. W. Brock, G. MacPherson, M. W. Reed, F. J. Couch, E. L. Goode, J. E. Olson, H. Meijers-Heijboer, A. van den Ouweland, F. Rivadeneira, R. L. Milne, G. Ribas, A. Gonzalez-Neira, J. Benitez, J. L. Hopper, M. McCredie, M. Southey, G. G. Giles, C. Schroen, C. Justenhoven, H. Brauch, U. Hamann, Y. D. Ko, A. B. Spurdle, J. Beesley, X. Chen, A. Mannermaa, V. M. Kosma, V. Kataja, J. Hartikainen, N. E. Day, et al.** 2007. Genome-wide association study identifies novel breast cancer susceptibility loci. Nature **447:**1087–1093.

10. **Ghoussaini, M., H. Song, T. Koessler, A. A. Al Olama, Z. Kote-Jarai, K. E. Driver, K. A. Pooley, S. J. Ramus, S. K. Kjaer, E. Hogdall, R. A. DiCioccio, A. S. Whittemore, G. A. Gayther, G. G. Giles, M. Guy, S. M. Edwards, J. Morrison, J. L. Donovan, F. C. Hamdy, D. P. Dearnaley, A. T. Ardern-Jones, A. L. Hall, L. T. O'Brien, B. N. Gehr-Swain, R. A. Wilkinson, P. M. Brown, J. L. Hopper, D. E. Neal, P. D. Pharoah, B. A. Ponder, R. A. Eeles, D. F. Easton, and A. M. Dunning.** 2008. Multiple loci with different cancer specificities within the 8q24 gene desert. J. Natl. Cancer Inst. **100:**962–966.

11. **Giles, R. H., J. H. van Es, and H. Clevers.** 2003. Caught up in a Wnt storm: Wnt signaling in cancer. Biochim. Biophys. Acta **1653:**1–24.

12. **Gimelbrant, A., J. N. Hutchinson, B. R. Thompson, and A. Chess.** 2007. Widespread monoallelic expression on human autosomes. Science **318:**1136–1140.

13. **Graham, M., J. M. Adams, and S. Cory.** 1985. Murine T lymphomas with retroviral inserts in the chromosomal 15 locus for plasmacytoma variant translocations. Nature **314:**740–743.

14. **Gustavson, M. D., H. C. Crawford, B. Fingleton, and L. M. Matrisian.** 2004. Tcf binding sequence and position determines beta-catenin and Lef-1 responsiveness of MMP-7 promoters. Mol. Carcinog. **41:**125–139.

15. **Haiman, C. A., N. Patterson, M. L. Freedman, S. R. Myers, M. C. Pike, A. Waliszewska, J. Neubauer, A. Tandon, C. Schirmer, G. J. McDonald, S. C. Greenway, D. O. Stram, L. Le Marchand, L. N. Kolonel, M. Frasco, D. Wong, L. C. Pooler, K. Ardlie, I. Oakley-Girvan, A. S. Whittemore, K. A. Cooney, E. M. John, S. A. Ingles, D. Altshuler, B. E. Henderson, and D. Reich.** 2007. Multiple regions within 8q24 independently affect risk for prostate cancer. Nat. Genet. **39:**638–644.

16. **He, T. C., A. B. Sparks, C. Rago, H. Hermeking, L. Zawel, L. T. da Costa, P. J. Morin, B. Vogelstein, and K. W. Kinzler.** 1998. Identification of c-MYC as a target of the APC pathway. Science **281:**1509–1512.

17. **Heintzman, N. D., G. C. Hon, R. D. Hawkins, P. Kheradpour, A. Stark, L. F. Harp, Z. Ye, L. K. Lee, R. K. Stuart, C. W. Ching, K. A. Ching, J. E. Antosiewicz-Bourget, H. Liu, X. Zhang, R. D. Green, V. V. Lobanenkov, R. Stewart, J. A. Thomson, G. E. Crawford, M. Kellis, and B. Ren.** 2009. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature **459:**108–112.

18. **Heintzman, N. D., R. K. Stuart, G. Hon, Y. Fu, C. W. Ching, R. D. Hawkins, L. O. Barrera, S. Van Calcar, C. Qu, K. A. Ching, W. Wang, Z. Weng, R. D. Green, G. E. Crawford, and B. Ren.** 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat. Genet. **39:**311–318.

19. **Henglein, B., H. Synovzik, P. Groitl, G. W. Bornkamm, P. Hartl, and M. Lipp.** 1989. Three breakpoints of variant t(2;8) translocations in Burkitt's lymphoma cells fall within a region 140 kilobases distal from c-myc. Mol. Cell. Biol. **9:**2105–2113.

20. **Kiemeney, L. A., S. Thorlacius, P. Sulem, F. Geller, K. K. Aben, S. N. Stacey, J. Gudmundsson, M. Jakobsdottir, J. T. Bergthorsson, A. Sigurdsson, T. Blondal, J. A. Witjes, S. H. Vermeulen, C. A. Hulsbergen-van de Kaa, D. W. Swinkels, M. Ploeg, E. B. Cornel, H. Vergunst, T. E. Thorgeirsson, D. Gudbjartsson, S. A. Gudjonsson, G. Thorleifsson, K. T. Kristinsson, M. Mouy, S. Snorradottir, D. Placidi, M. Campagna, C. Arici, K. Koppova, E. Gurzau, P. Rudnai, E. Kellen, S. Polidoro, S. Guarrera, C. Sacerdote, M. Sanchez, B. Saez, G. Valdivia, C. Ryk, P. de Verdier, A. Lindblom, K. Golka, D. T. Bishop, M. A. Knowles, S. Nikulasson, V. Petursdottir, E. Jonsson, G. Geirsson, B. Kristjansson, J. I. Mayordomo, G. Steineck, S. Porru, F. Buntinx, M. P. Zeegers, T. Fletcher, R. Kumar, G. Matullo, P. Vineis, A. E. Kiltie, J. R. Gulcher, U. Thorsteinsdottir, A. Kong, T. Rafnar, and K. Stefansson.** 2008. Sequence variant on 8q24 confers susceptibility to urinary bladder cancer. Nat. Genet. **40:**1307–1312.

21. **Kinzler, K. W., and B. Vogelstein.** 1996. Lessons from hereditary colorectal cancer. Cell **87:**159–170.

22. **Korinek, V., N. Barker, P. J. Morin, D. van Wichen, R. de Weger, K. W. Kinzler, B. Vogelstein, and H. Clevers.** 1997. Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC$^{-/-}$ colon carcinoma. Science **275:**1784–1787.

23. **Kuehl, W. M., L. A. Brents, M. Chesi, and P. L. Bergsagel.** 1996. Selective expression of one c-myc allele in two human myeloma cell lines. Cancer Res. **56:**4370–4373.

24. **Lieberman-Aiden, E., N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, and J. Dekker.** 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science **326:**289–293.

25. **Liu, J., and D. Levens.** 2006. Making myc. Curr. Top. Microbiol. Immunol. **302:**1–32.

26. **Lupien, M., J. Eeckhoute, C. A. Meyer, Q. Wang, Y. Zhang, W. Li, J. S. Carroll, X. S. Liu, and M. Brown.** 2008. FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. Cell **132:**958–970.

27. **Miele, A., and J. Dekker.** 2008. Long-range chromosomal interactions and gene regulation. Mol. Biosyst. **4:**1046–1057.

28. **Miele, A., N. Gheldof, T. M. Tabuchi, J. Dostie, and J. Dekker.** 2006. Unit 21.11, Mapping chromatin interactions by chromosome conformation capture. Curr. Protoc. Mol. Biol. doi:10.1002/0471142727.mb2111s74.

29. **Morin, P. J., A. B. Sparks, V. Korinek, N. Barker, H. Clevers, B. Vogelstein, and K. W. Kinzler.** 1997. Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. Science **275:**1787–1790.

30. **Nesbit, C. E., J. M. Tersak, and E. V. Prochownik.** 1999. MYC oncogenes and human neoplastic disease. Oncogene **18:**3004–3016.

31. **Nobrega, M. A., I. Ovcharenko, V. Afzal, and E. M. Rubin.** 2003. Scanning human gene deserts for long-range enhancers. Science **302:**413.

32. **Nusse, R.** 2005. Wnt signaling in disease and in development. Cell Res. **15:**28–32.

33. **Pomerantz, M. M., N. Ahmadiyeh, L. Jia, P. Herman, M. P. Verzi, H. Doddapaneni, C. A. Beckwith, J. A. Chan, A. Hills, M. Davis, K. Yao, S. M. Kehoe, H. J. Lenz, C. A. Haiman, C. Yan, B. E. Henderson, B. Frenkel, J. Barretina, A. Bass, J. Tabernero, J. Baselga, M. M. Regan, J. R. Manak, R. Shivdasani, G. A. Coetzee, and M. L. Freedman.** 2009. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. Nat. Genet. **41:**882–884.

34. **Pomerantz, M. M., C. A. Beckwith, M. M. Regan, S. K. Wyman, G. Petrovics, Y. Chen, D. J. Hawksworth, F. R. Schumacher, L. Mucci, K. L. Penney, M. J. Stampfer, J. A. Chan, K. G. Ardlie, B. R. Fritz, R. K. Parkin, D. W. Lin, M. Dyke, P. Herman, S. Lee, W. K. Oh, P. W. Kantoff, M. Tewari, D. G. McLeod, S. Srivastava, and M. L. Freedman.** 2009. Evaluation of the 8q24 prostate cancer risk locus and MYC expression. Cancer Res. **69:**5568–5574.

35. **Salinas, C. A., E. Kwon, C. S. Carlson, J. S. Koopmeiners, Z. Feng, D. M. Karyadi, E. A. Ostrander, and J. L. Stanford.** 2008. Multiple independent genetic variants in the 8q24 region are associated with prostate cancer risk. Cancer Epidemiol. Biomarkers Prev. **17:**1203–1213.

36. **Sansom, O. J., V. S. Meniel, V. Muncan, T. J. Phesse, J. A. Wilkins, K. R. Reed, J. K. Vass, D. Athineos, H. Clevers, and A. R. Clarke.** 2007. Myc deletion rescues Apc deficiency in the small intestine. Nature **446:**676–679.

37. **Shachaf, C. M., A. J. Gentles, S. Elchuri, D. Sahoo, Y. Soen, O. Sharpe, O. D. Perez, M. Chang, D. Mitchel, W. H. Robinson, D. Dill, G. P. Nolan, S. K. Plevritis, and D. W. Felsher.** 2008. Genomic and proteomic analysis reveals a threshold level of MYC required for tumor maintenance. Cancer Res. **68:**5132–5142.

38. **Tetsu, O., and F. McCormick.** 1999. Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. Nature **398:**422–426.

39. **Tolhuis, B., R. J. Palstra, E. Splinter, F. Grosveld, and W. de Laat.** 2002. Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol. Cell **10:**1453–1465.

40. **Tomlinson, I., E. Webb, L. Carvajal-Carmona, P. Broderick, Z. Kemp, S. Spain, S. Penegar, I. Chandler, M. Gorman, W. Wood, E. Barclay, S. Lubbe, L. Martin, G. Sellick, E. Jaeger, R. Hubner, R. Wild, A. Rowan, S. Fielding, K. Howarth, A. Silver, W. Atkin, K. Muir, R. Logan, D. Kerr, E. Johnstone, O. Sieber, R. Gray, H. Thomas, J. Peto, J. B. Cazier, and R. Houlston.** 2007. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. Nat. Genet. **39:**984–988.

41. **Tsichlis, P. N., B. M. Shepherd, and S. E. Bear.** 1989. Activation of the Mlvi-1/mis1/pvt-1 locus in Moloney murine leukemia virus-induced T-cell lymphomas. Proc. Natl. Acad. Sci. U. S. A. **86:**5487–5491.

42. **Tuupanen, S., M. Turunen, R. Lehtonen, O. Hallikas, S. Vanharanta, T. Kivioja, M. Bjorklund, G. Wei, J. Yan, I. Niittymaki, J. P. Mecklin, H. Jarvinen, A. Ristimaki, M. Di-Bernardo, P. East, L. Carvajal-Carmona, R. S. Houlston, I. Tomlinson, K. Palin, E. Ukkonen, A. Karhu, J. Taipale, and L. A. Aaltonen.** 2009. The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling. Nat. Genet. **41:**885–890.

43. **van de Wetering, M., E. Sancho, C. Verweij, W. de Lau, I. Oving, A. Hurl-**

stone, K. van der Horn, E. Batlle, D. Coudreuse, A. P. Haramis, M. Tjon-Pon-Fong, P. Moerer, M. van den Born, G. Soete, S. Pals, M. Eilers, R. Medema, and H. Clevers. 2002. The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. Cell **111:**241–250.

44. Wingender, E., X. Chen, R. Hehl, H. Karas, I. Liebich, V. Matys, T. Meinhardt, M. Pruss, I. Reuter, and F. Schacherer. 2000. TRANSFAC: an integrated system for gene expression regulation. Nucleic Acids Res. **28:**316–319.

45. Yekkala, K., and T. A. Baudino. 2007. Inhibition of intestinal polyposis with

reduced angiogenesis in ApcMin/+ mice due to decreases in c-Myc expression. Mol. Cancer Res. **5:**1296–1303.

46. Yochum, G. S., R. Cleland, and R. H. Goodman. 2008. A genome-wide screen for β-catenin binding sites identifies a downstream enhancer element that controls c-*Myc* gene expression. Mol. Cell. Biol. **28:**7368–7379.

47. Yochum, G. S., C. M. Sherrick, M. Macpartlin, and R. H. Goodman. 2010. A β-catenin/TCF-coordinated chromatin loop at MYC integrates 5′ and 3′ Wnt responsive enhancers. Proc. Natl. Acad. Sci. U. S. A. **107:** 145–150.